SEQUENCE LISTING

<110> ALVES, ALEXANDRA M.C.R.
RECORD, ERIC
LOMASCOLO, ANNE
SIGOILLOT, JEAN-CLAUDE
ASTHER, MARCEL
WOSTEN, HAN A.B.

<120> METHOD FOR OVERPRODUCING A SPECIFIC RECOMBINANT PROTEIN WITH P. CINNABARINUS MONOKARYOTIC STRAINS

<130> 0508-1167

<140> 10/586,348
<141> 2006-07-14

<150> PCT/FR05/000093
<151> 2005-01-14

<150> FR 04/00366
<151> 2004-01-15

<160> 18

<170> PatentIn Ver. 3.3

<210> 1
<211> 3330
<212> DNA
<213> Pycnoporus cinnabarinus

<220>
<221> CDS
<222> (128)..(310)

<220>
<221> CDS
<222> (368)..(436)

<220>
<221> CDS
<222> (490)..(610)

<220>
<221> CDS
<222> (664)..(777)

<220>
<221> CDS
<222> (833)..(896)

<220>
<221> CDS
<222> (960)..(1055)

<220>
<221> CDS
<222> (1114)..(1270)

<220>
<221> CDS
<222> (1334)..(1531)

<220>
<221> CDS
<222> (1592)..(1648)

<220>
<221> CDS
<222> (1705)..(1911)

<220>
<221> CDS
<222> (1968)..(2255)

<400> 1

```
ctgcagacat ctggagcgcc tgtctttccc ctagtataaa tgatgtctgt ccgcaggtcc     60
ttgaagaccg ctcgagtccc acttgagttt taggtaggac ctgtccacca aacccctctt    120
tctgatc atg tcg agg ttc cag tcc ctc ttc ttc ttc gtc ctc gtc tcc      169
        Met Ser Arg Phe Gln Ser Leu Phe Phe Phe Val Leu Val Ser
        1               5                   10

ctc acc gct gtg gcc aac gca gcc ata ggg cct gtg gcg gac ctg acc      217
Leu Thr Ala Val Ala Asn Ala Ala Ile Gly Pro Val Ala Asp Leu Thr
15                  20                  25                  30

ctt acc aat gcc cag gtc agc ccc gat ggc ttc gct cgc gag gcc gtc      265
Leu Thr Asn Ala Gln Val Ser Pro Asp Gly Phe Ala Arg Glu Ala Val
                35                  40                  45

gtg gtg aac ggt atc acc cct gcc cct ctc atc aca ggc aat aag          310
Val Val Asn Gly Ile Thr Pro Ala Pro Leu Ile Thr Gly Asn Lys
            50                  55                  60

gtatgtatat gctgctcgtc cctcagagct acatacatct gatccacaat cgtttag       367

ggc gat cga ttc cag ctc aat gtc atc gac cag ttg aca aat cat acc      415
Gly Asp Arg Phe Gln Leu Asn Val Ile Asp Gln Leu Thr Asn His Thr
            65                  70                  75

atg ttg aaa aca tct agt att gtaagggttc agtttttccc gactaccatg         466
Met Leu Lys Thr Ser Ser Ile
        80

ttattgacca tcaccactcg tag cat tgg cac ggc ttc ttc cag caa ggc acg    519
                        His Trp His Gly Phe Phe Gln Gln Gly Thr
                        85                  90

aac tgg gcc gat ggt ccc gcg ttc gtg aac cag tgt ccc atc gct tcg      567
Asn Trp Ala Asp Gly Pro Ala Phe Val Asn Gln Cys Pro Ile Ala Ser
95                  100                 105                 110
```

```
ggc cac tcg ttc ttg tat gac ttt caa gtt ccc gac caa gca g                610
Gly His Ser Phe Leu Tyr Asp Phe Gln Val Pro Asp Gln Ala
                115                 120

gtacgaattc cgtacacgtt tcattgcgtc gcaactaaac ctcctcttac tag gg             665
                                                          Gly
                                                          125

act ttc tgg tac cat agc cat ctc tcc acg caa tac tgc gat ggt ttg           713
Thr Phe Trp Tyr His Ser His Leu Ser Thr Gln Tyr Cys Asp Gly Leu
                130                 135                 140

agg ggg cct ttc gtc gtc tac gac ccc aac gat cct cac gct agc ctg           761
Arg Gly Pro Phe Val Val Tyr Asp Pro Asn Asp Pro His Ala Ser Leu
            145                 150                 155

tat gac att gat aac g gtgagcagat catggtatcg caatattgcg tccacttatg        817
Tyr Asp Ile Asp Asn
        160

cttcctggca tccag ac  gac act gtc att acg ctg gct gat tgg tat cac          867
                 Asp Asp Thr Val Ile Thr Leu Ala Asp Trp Tyr His
                         165                 170

gtt gct gcc aag ctc gga cct cgc ttc cc  gtacgtgtca aatgtctacg             916
Val Ala Ala Lys Leu Gly Pro Arg Phe Pro
175                 180

agagatctca catatacgac tagactcact tcgctgatta cag a ttt ggc tcc gat         972
                                                  Phe Gly Ser Asp
                                                  185

tca acc ctt atc aat gga ctt ggt cga acc act ggc ata gca ccg tcc           1020
Ser Thr Leu Ile Asn Gly Leu Gly Arg Thr Thr Gly Ile Ala Pro Ser
    190                 195                 200

gac ttg gca gtt atc aag gtc acg cag ggc aag cg  gtaagtatgg               1065
Asp Leu Ala Val Ile Lys Val Thr Gln Gly Lys Arg
205                 210                 215

atggtcatca ctgcacattg gctctgatac atggccttgt ttccacag c tac cgc           1120
                                                        Tyr Arg

ttc cgc ttg gtg tcg ctt tct tgc gat ccg aac cat aca ttc agc att           1168
Phe Arg Leu Val Ser Leu Ser Cys Asp Pro Asn His Thr Phe Ser Ile
    220                 225                 230

gat aat cac aca atg act ata att gag gcg gac tcg atc aac act caa           1216
Asp Asn His Thr Met Thr Ile Ile Glu Ala Asp Ser Ile Asn Thr Gln
235                 240                 245                 250

ccc cta gag gtt gat tca atc cag att ttt gcc gcg cag cgc tac tcc           1264
Pro Leu Glu Val Asp Ser Ile Gln Ile Phe Ala Ala Gln Arg Tyr Ser
                255                 260                 265

ttc gtg gtaggtcgta ggctcctgtc atcaagtttg cagacattct tagatacacc           1320
Phe Val
```

```
tttttcaatg cag ctg gat gct agc cag ccg gtg gat aac tac tgg atc      1369
               Leu Asp Ala Ser Gln Pro Val Asp Asn Tyr Trp Ile
                   270                 275                 280

cgc gca aac cct gcc ttc gga aac aca ggt ttt gct ggt gga atc aat      1417
Arg Ala Asn Pro Ala Phe Gly Asn Thr Gly Phe Ala Gly Gly Ile Asn
               285                 290                 295

tct gcc atc ctg cgt tat gat ggc gca ccc gag atc gag cct acg tct      1465
Ser Ala Ile Leu Arg Tyr Asp Gly Ala Pro Glu Ile Glu Pro Thr Ser
           300                 305                 310

gtc cag act act cct acg aag cct ctg aac gag gtc gac ttg cat cct      1513
Val Gln Thr Thr Pro Thr Lys Pro Leu Asn Glu Val Asp Leu His Pro
       315                 320                 325

ctc tcg cct atg cct gtg gtacgtgtct caaagaacct cgatcactaa            1561
Leu Ser Pro Met Pro Val
    330

gtgcatgtca actcatatgg tgcatgacag cct ggc agc ccc gag ccc gga ggt     1615
                                Pro Gly Ser Pro Glu Pro Gly Gly
                                335                 340

gtc gac aag cct ctg aac ttg gtc ttc aac ttc gtgagtactg gcgcgcttcc    1668
Val Asp Lys Pro Leu Asn Leu Val Phe Asn Phe
        345                 350

gtagcacacg ttcgaacaaa gcctgatacc atgcag aac ggc acc aac ttc ttc      1722
                                        Asn Gly Thr Asn Phe Phe
                                            355

atc aac gac cac acc ttt gtc ccg ccg tct gtc cca gtc ttg cta caa      1770
Ile Asn Asp His Thr Phe Val Pro Pro Ser Val Pro Val Leu Leu Gln
360                 365                 370                 375

atc ctc agt ggg gcg cag gcg gct cag gac ctg gtc ccg gag ggc agc      1818
Ile Leu Ser Gly Ala Gln Ala Ala Gln Asp Leu Val Pro Glu Gly Ser
                380                 385                 390

gtg ttc gtt ctt ccc agc aac tcg tcc att gag ata tcc ttc cct gcc      1866
Val Phe Val Leu Pro Ser Asn Ser Ser Ile Glu Ile Ser Phe Pro Ala
            395                 400                 405

act gcc aat gcc cct gga ttc ccc cat ccg ttc cac ttg cac ggt         1911
Thr Ala Asn Ala Pro Gly Phe Pro His Pro Phe His Leu His Gly
        410                 415                 420

gtacgtctgc cttcccctcg tctaaaggcg gagtcgatat ctgactccca tcacag cac   1970
                                                            His

gcc ttc gct gtc gtc cgg agc gcc ggg agc agc gtc tac aac tac gac      2018
Ala Phe Ala Val Val Arg Ser Ala Gly Ser Ser Val Tyr Asn Tyr Asp
    425                 430                 435
```

```
aac ccg atc ttc cgc gac gtc gtc agc acc ggc cag ccc ggc gac aac     2066
Asn Pro Ile Phe Arg Asp Val Val Ser Thr Gly Gln Pro Gly Asp Asn
440                 445                 450                 455

gtc acg att cgc ttc gag acc aat aac cca ggc ccg tgg ttc ctc cac     2114
Val Thr Ile Arg Phe Glu Thr Asn Asn Pro Gly Pro Trp Phe Leu His
                460                 465                 470

tgc cac att gac ttc cac ctc gac gca ggc ttt gct gta gtc atg gcc     2162
Cys His Ile Asp Phe His Leu Asp Ala Gly Phe Ala Val Val Met Ala
            475                 480                 485

gag gac act ccg gac acc aag gcc gcg aac cct gtt cct cag gcg tgg     2210
Glu Asp Thr Pro Asp Thr Lys Ala Ala Asn Pro Val Pro Gln Ala Trp
        490                 495                 500

tcg gac ttg tgc ccc atc tat gat gca ctt gac ccc agc gac ctc         2255
Ser Asp Leu Cys Pro Ile Tyr Asp Ala Leu Asp Pro Ser Asp Leu
    505                 510                 515

tgagcgggat tgttactgtg acctggtgtg gggggaacat gtcgagggct ttcatcgatc   2315
agggactttc aaggttggca taatatacct cacggcctgg atgactcgga cagcgtgtgg   2375
gcgtgggtgt aactctgctt gatgttgaaa aaaggatttt atgtagaaca atttatgagc   2435
aatcagcaat caataggatt gtgtcggttt cgacgaaatg tcttgtctcc ctgacattac   2495
ttttgtgcga gaaatgggtc catgatacac atcattgagc tctcaatacc aagaaggatt   2555
acccatgtca atacccaaga tcatgtcttc gctgtccgca atggtctcat gttgcgttga   2615
gcagatcgca gtacgttgaa aagcgattag tattacatgc aacatgcaac atttggaagg   2675
gggcatgcag aggttcagct cgcgtcagtc ggccaagtag cgacctttgc cgcactgcct   2735
gttaacctga acgtatgctt cagaactccg tcggtatcga gagcgatcgt gtacgttccg   2795
ggatagatcc attgatcccc gctctggtcg gcgcgtgcga tggccccgag cgtcaccggc   2855
agcttcgcga tcgcgctttt cctaggggcg aggccgtgta cccgcgtgta cgagacgagc   2915
tgcttgttcg ggtggggcga aggcccgaag gagccactca cgaagagcaa tgcgacgtaa   2975
tccgaggtag ccttgcccgt gttagtcaca cgcacggaga acgtgtcgag cggcgcgagg   3035
tcgaggaagg cggcgctctt ctgaccgcgc tgtacgaggt cggaaatcga atacgtcgat   3095
ggcggtcctc caaagtccgt gacgttggtc gcatcggccg ccgcgcctgg agctgcccaa   3155
gagaaatcga aggtggtgaa gtgcagtcca aagccaaatt cgtagaccgg cgtgccggtg   3215
taccacttgt atgtacgccc cgggttcgac gcgcttgggc gaagggtcat gtcagtcatc   3275
ggaacctgat cagcgtagat ggctgggtat tgggtgatgg gcaggcgtcc tgcag        3330
```

<210> 2
<211> 518
<212> PRT
<213> Pycnoporus cinnabarinus

<400> 2

```
Met Ser Arg Phe Gln Ser Leu Phe Phe Phe Val Leu Val Ser Leu Thr
1               5                  10                  15

Ala Val Ala Asn Ala Ala Ile Gly Pro Val Ala Asp Leu Thr Leu Thr
            20                  25                  30

Asn Ala Gln Val Ser Pro Asp Gly Phe Ala Arg Glu Ala Val Val Val
        35                  40                  45

Asn Gly Ile Thr Pro Ala Pro Leu Ile Thr Gly Asn Lys Gly Asp Arg
    50                  55                  60
```

Phe Gln Leu Asn Val Ile Asp Gln Leu Thr Asn His Thr Met Leu Lys
65 70 75 80

Thr Ser Ser Ile His Trp His Gly Phe Phe Gln Gln Gly Thr Asn Trp
85 90 95

Ala Asp Gly Pro Ala Phe Val Asn Gln Cys Pro Ile Ala Ser Gly His
100 105 110

Ser Phe Leu Tyr Asp Phe Gln Val Pro Asp Gln Ala Gly Thr Phe Trp
115 120 125

Tyr His Ser His Leu Ser Thr Gln Tyr Cys Asp Gly Leu Arg Gly Pro
130 135 140

Phe Val Val Tyr Asp Pro Asn Asp Pro His Ala Ser Leu Tyr Asp Ile
145 150 155 160

Asp Asn Asp Asp Thr Val Ile Thr Leu Ala Asp Trp Tyr His Val Ala
165 170 175

Ala Lys Leu Gly Pro Arg Phe Pro Phe Gly Ser Asp Ser Thr Leu Ile
180 185 190

Asn Gly Leu Gly Arg Thr Thr Gly Ile Ala Pro Ser Asp Leu Ala Val
195 200 205

Ile Lys Val Thr Gln Gly Lys Arg Tyr Arg Phe Arg Leu Val Ser Leu
210 215 220

Ser Cys Asp Pro Asn His Thr Phe Ser Ile Asp Asn His Thr Met Thr
225 230 235 240

Ile Ile Glu Ala Asp Ser Ile Asn Thr Gln Pro Leu Glu Val Asp Ser
245 250 255

Ile Gln Ile Phe Ala Ala Gln Arg Tyr Ser Phe Val Leu Asp Ala Ser
260 265 270

Gln Pro Val Asp Asn Tyr Trp Ile Arg Ala Asn Pro Ala Phe Gly Asn
275 280 285

Thr Gly Phe Ala Gly Gly Ile Asn Ser Ala Ile Leu Arg Tyr Asp Gly
290 295 300

Ala Pro Glu Ile Glu Pro Thr Ser Val Gln Thr Thr Pro Thr Lys Pro
305 310 315 320

Leu Asn Glu Val Asp Leu His Pro Leu Ser Pro Met Pro Val Pro Gly
325 330 335

Ser Pro Glu Pro Gly Gly Val Asp Lys Pro Leu Asn Leu Val Phe Asn
340 345 350

Phe Asn Gly Thr Asn Phe Phe Ile Asn Asp His Thr Phe Val Pro Pro
355 360 365

```
Ser Val Pro Val Leu Leu Gln Ile Leu Ser Gly Ala Gln Ala Ala Gln
    370                 375                 380

Asp Leu Val Pro Glu Gly Ser Val Phe Val Leu Pro Ser Asn Ser Ser
385                 390                 395                 400

Ile Glu Ile Ser Phe Pro Ala Thr Ala Asn Ala Pro Gly Phe Pro His
                405                 410                 415

Pro Phe His Leu His Gly His Ala Phe Ala Val Val Arg Ser Ala Gly
            420                 425                 430

Ser Ser Val Tyr Asn Tyr Asp Asn Pro Ile Phe Arg Asp Val Val Ser
        435                 440                 445

Thr Gly Gln Pro Gly Asp Asn Val Thr Ile Arg Phe Glu Thr Asn Asn
    450                 455                 460

Pro Gly Pro Trp Phe Leu His Cys His Ile Asp Phe His Leu Asp Ala
465                 470                 475                 480

Gly Phe Ala Val Val Met Ala Glu Asp Thr Pro Asp Thr Lys Ala Ala
                485                 490                 495

Asn Pro Val Pro Gln Ala Trp Ser Asp Leu Cys Pro Ile Tyr Asp Ala
            500                 505                 510

Leu Asp Pro Ser Asp Leu
        515
```

<210> 3
<211> 2527
<212> DNA
<213> Pycnoporus cinnabarinus

<400> 3

```
agatctccga accagaaatg cgattgcgtt caggcccaat taagaataaa gctgcgtcag 60
ggcagcgacg tatcttgatc catcattgac tcaccggcat cggcgtcaac accaaagcaa 120
gctcgtccca cccataggcg tgcaccggcc ggcgtgcgcc attgaggtac atgagcgggg 180
cgaaagtccg ccattggtag ccctgtcgtg gacgcgcggc gatgaaacgt ttcccaccat 240
tgggaagaaa cgtctgcggc ccatcatccc ttcaccggat gacaaggcgg cgtcgcgcct 300
ttgccgcaga ggccggcggg cgacatgcac agcgaaggtc cgttgcggat gggaagcagg 360
caatcagtgg gtgtcctacg ccgccacgat ggtcggggag cgtaggcgcc ctcccataag 420
gcggcaagca tcatgatgct ctccgattcg ggaagcctgg tgcgatgctg gagagactct 480
ctccgagaga ccagtgtgcg caacgttcct ggcctggaag actttaaagt gagtgtagaa 540
gggcgagcag aggacgatca tcggattgca ggaacccatcg gcatcctcag cctgggaagg 600
atggctcttg gtagacattc gcggaaggtg tcctagatgt gagcgggctt cttggatgat 660
catgtcgtaa ctttttctga cctcgtcggt ggtacgcatg gcaggattga gcattacggt 720
atgcctccca ttcataaacg ataaccccctt ccttcaggtt ggtcatctcc atagagcggc 780
acgctctcaa ggcctaggct attcacacct ccttcgcaac atccctattc acggtgtctg 840
taaggaacga cttgtcatgg gatcacatga agtgcagcat actgttcgcc ggtctcgcag 900
tacagacgct agtacgggaa gtcgacatcc aagcgttcag tcaccacatg gcaaaaaagc 960
tgcaccatac tctttatggt gagttgttcg tgagtggtat acagtcattc atgagggaat 1020
gcccaccgga tagggtgtgg cggccgcaat attcatcgcc tggcaatagt cgatgtgcgt 1080
ccttgttcaa tgaatatcat gggtcacatg tggagacggt taaacagcgt tgactgtgaa 1140
tccctggtgt gtgttgggcc gaacaggtac gttgcaggaa caccaatatc tcttcggcag 1200
```

cccagttctt tgcgagcggc acaggcaggc atcgcgcaac agatcccagc catccggcct 1260
ctgacattcg ggatacctga agcccttcag gtacggagcg aagaggtggg ctctctgcag 1320
cgattggcgg acggatagct gtatttcctc tctcaccatt gggaagatgt gaaaggctcc 1380
atcatatagc ggctcaactc tacctcgaat gtccaaacac ggcgggaata cttatttatg 1440
tggacaaggc cgagctatga tagcttgctc ccgaagttgg taagtcccgc aatctgcggt 1500
tcaggcaaca gtctcggaaa aataagaaga atattgtagg tgcgtgtagg cgtatcgccc 1560
aaatgcgcac acacggaggc tttaggagat gaagcgcccg tgagcggtaa gggagttggt 1620
tcaccgccgc cccgaccgac tctctctctt tcccagcatc atgtctcggc gcaaacttta 1680
ccctctattg accaactcca cgagaaagca ggaacagctt ccttgtctct catgacgtcc 1740
gcaatccaga cccttagccg gttcgttact catcgttatc cctgccgcca tggtagtgga 1800
gtcagcctgg ccagtgcgta gtcccgtctc tcttgctgca ctagagaagc cccatgagac 1860
agcgtttttt gctttatttc tgctgtttct atagacacca taggggcaaa cgatcctgca 1920
cgcccagagg tattgggctc gtcagattcc cagttttttct cctcggtctg aatcggctgc 1980
acggcagata aatcggccgg aaatgctata gcccttcata gcccgctatg agagtcgcaa 2040
aaggcttgtc agtcaggtcg gtcgagtggc tctcacgaag agcgtcaact tcgcgcgaca 2100
gccgcctttc agggcaagat agatcctccc atcatcccct actgcgctca gcgccggtac 2160
cgaacaattg acttaccgac atcctccggg acgcgcaaat gctgttcgac ggaacgtaat 2220
cctcttcgtc ccgcctcttt tcgctctcac gcattccgtg tggttcgcgc gacggccgct 2280
catcaggacc agaccagtct caatgtctgg taccggcaca atggtgacac tgcggcaact 2340
gagtaggtct ggtcactctg gtgcaccgtc gcttacgctg accttcggga tactgtcctg 2400
cagacatctg gagcgcctgt ctttccccta gtataaatga tgtctgtccg caggtccttg 2460
aagaccgctc gagtcccact tgagttttag gtaggacctg tccaccaaac ccctctttct 2520
gatcatg 2527

<210> 4
<211> 643
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic vector

<400> 4
cgaccgagcg cgcgccaccc agcctatccc gcgcgggtcg ggacccaaaa taagcgggcc 60
ccgccgcgcc ccgtcgggcg agcgggtgta tctacgaacg gaactgggag gcgactcgga 120
agagtttggt tagaaagggg aacaccatcg cggacggccc agtgctctgg dcagctgagc 180
gtgcattgtg ttcaattctg acctgtggca tgtaaggaac gtgctcggga tcggagggtg 240
gcgcgagagc ctcttcggtg tgagattagt aactgtactg cgaagccgcg gaggggttag 300
gatgagaggt agacagggtc gcagcccagg tgcgagaagg actgcgaagg actgttcttc 360
gaccgcgcac ctgcaattgc gcgcatggat agaatagagc gtcgccctcg aggggactc 420
gaccagggct ggtggtggcg cccgacggga ctggctgggc atttgcagat ggcgcgcagt 480
ccaggccgcc gccgatgtgt tcatcccgtt ttgtcagtat cgatcggatc tttcgggcgt 540
gggtataaaa gcgcgccgcc cgccgtctcc ctctttctcc agcactccca tccagagcac 600
ttccctctcc catcgcatcc catcacacaa taatgcccat cac 643

<210> 5
<211> 1033
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic vector

<400> 5
agcttctccg gccccgaatc gaacggcagg atgtgtgggc gtgtccaata ttgccatgaa 60
aatctgtcag aagtgagccc tctcgtcacc ctgtacagct tcgctgagtt gaaaagcagg 120
gttcatcttg ggctcactga tgcactgagc tcgaccggag aactaaatga ccagccggag 180
tgttcactaa cttaacgccg ggtattcagg gcagcttctc tatgttgcgc ctacgacgta 240
gatcaccgcc catgaacggg ggaaacgggg aggggtgcgt ttggtacgtc tttacgtctg 300
gctatgttgt attgaccagc gtctgcagaa gatgggcacg acgatgcgcc gagccggcca 360
gtgtcgtcgg atgtccactg ttgaggccat ccttttgcta gacagacgga agagctttgg 420
aggtgcgatt cctctacgaa tgggaagggg cttagatgga gagtgacacg tctgagctcc 480
ccaacacgcc ttcgccgagg gtgcgtctcc gcggacattc acctcagttc attgttctga 540
cctgcctaat tgtatagacc ggccaacaac cttgctgacg cccatcataa cagtgccctg 600
cacagagcct tcccactcag tcggcgcctc cctcaatcaa tcccactaac tcgccggctc 660
tgcccccttcg ccgctcgaca cgtcgcttgg aagagcccgg gcacgggcgt ccgctccccc 720
cttccctccg cgtcgtcatg cacgcagcgt taatgttgct gcaggcgagc cgtaagtata 780
ttcaaaggcg tagcgaatga atagcaggcg cgcggggacc tggcacgcgc ggcatgaaca 840
tgcagacttg ggtgacgata acttgaactc agacgcggcg aatgaatatc caaacgcgcg 900
ggaagaaaat aatttacggg agcctcccca ggtataaaag cccctcaccc gctcactctt 960
tctccagtcg aacaccccag ttcaactacc cagcccttcc ttccttcgct atccttcytt 1020
acaacctgct cgc 1033

<210> 6
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 6
caytggcayg grttcttcc 19

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 7
gagrtggaag tcratgtgrc 20

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 8
ggataactac tggatccgcg 20

<210> 9
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 9
cgcagtattg cgtggagag 19

<210> 10
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 10
gacatctgga gcgcctgtc 19

<210> 11
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 11
atcgaaggtt ccgatgactg acatgac 27

<210> 12
<211> 5122
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic vector

<400> 12
catgggatat cgcatgcctg cagagctcta gagtcgacgg gcccggtacc gcggccgcct 60
taagacgcgt ggatccgcag gtgaacgcgc ctatcggtgg gatattcggg cgacgggagc 120
ctcggcaatc tgagcctcgt tactgcctag caaattcgga atcccttcga tgtcataggg 180
tcgcggacaa gtgatcgtct tgctacatac tccaaggtgt tgactcattc cctcgataat 240
gaacattgtt gttgttgttt gttctctatc cgctcagtca cgcgacccca cacgtgcatg 300
gttgaacttc gccacgcaac aaccgcatga cgacatggcg aacctaagta aaggctgagt 360
cgtggactaa agcactccac tttacggcga ggatgccagt ctacgtcatg aatgaagcct 420

caggtcccga agtaaggggg tacaaaagga gggtgaaagg tggacgtttt cttaccatcc 480
ttccacctcc cagaccacca tgccgggaat tcccagcttg ctcaaaaagg ttctgcccgt 540
acgcccgcga aattccttcg aggtggcccc tatcgcatac atgcacgact tcaaaacatc 600
cattctatca ttttgggatc gtacaattat tagacatgtt gtacaacgtt acattccttt 660
cttcttttac tctccggccc agtctatgta gaggtaaagt acaagcgtcc aaaggatcag 720
gcacttagag cgcgccgtct tgcttcgccg cttagagcgc gccgtcctgc ttcgccgcgt 780
agacgagcag gtcgcagaca cggcgggagt agcccccactc gttgtcgtac caggcaatga 840
gcttcacgaa gctcttgctg atcgcgatgc cggggatcga tccacgcgtc ttaaggcggc 900
cgcggtaccc cctcggaccc gtcgggccgc gtcggaccgg cggtgttggt cggcgtcggt 960
cagtcctgct cctcggccac gaagtgcacg cagttgccgg ccgggtcgcg cagggcgaac 1020
tcccgccccc acggctgctc gccgatctcg gtcatggccg gcccggaggc gtcccggaag 1080
ttcgtggaca cgacctccga ccactcggcg tacagctcgt ccaggccgcg cacccacacc 1140
caggccaggg tgttgtccgg caccacctgg tcctggaccg cgctgatgaa cagggtcacg 1200
tcgtcccgga ccacaccggc gaagtcgtcc tccacgaagt cccgggagaa cccgagccgg 1260
tcggtccaga actcgaccgc tccggcgacg tcgcgcgcgg tgagcaccgg aacggcactg 1320
gtcaacttgg ccatgcatgg tgatgggcat tatgtgtgat gggatgcgat gggagaggga 1380
agtgctctgg atgggagtgc tggagaaaga gggagacggc gggcggcgcg ccttttatac 1440
ccacgcccga aagatccgat cgatactgac aaaacgggat gaacacatcg gcggcggcct 1500
ggactgcgcg ccatctgcaa atgcccagcc agtcccgtcg ggcgccacca ccagccctgg 1560
tcgagtcccc ctcgagggcg acgctctatt ctatccatgc gcgcaattgc aggtgcgcgg 1620
tcgaagaaca gtccttcgca gtccttctcg cacctgggct gcgaccctgt ctacctctca 1680
tcctaacccc tccgcggctt cgcagtacag ttactaatct cacaccgaag aggctctcgc 1740
gccaccctcc gatcccgagc acgttcctta catgccacag cgtcagaatt gaacacaatg 1800
cacgtcarat cagatccccg ggaattcgta atcatggtca tagctgtttc ctgtgtgaaa 1860
ttgttatccg ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg 1920
gggtgcctaa tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca 1980
gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg 2040
tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg 2100
gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg 2160
ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 2220
ggccgcgttg ctggcgtttt tccataggct ccgccccct gacgagcatc acaaaaatcg 2280
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc 2340
tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc 2400
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc 2460
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg 2520
ctgcgccttа tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc 2580
actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 2640
gttcttgaag tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc 2700
tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac 2760
caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 2820
atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 2880
acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa 2940
ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta 3000
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 3060
tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag 3120
tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag caataaacca 3180
gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc 3240
tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt 3300
tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag 3360
ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt 3420
tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat 3480
ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt 3540
gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc 3600
ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat 3660
cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag 3720
ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt 3780
ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg 3840
gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta 3900

ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc 3960
gcgcacattt ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt 4020
aacctataaa aataggcgta tcacgaggcc ctttcgtctc gcgcgtttcg gtgatgacgg 4080
tgaaaacctc tgacacatgc agctcccgga gacggtcaca gcttgtctgt aagcggatgc 4140
cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcgggtgtc ggggctggct 4200
taactatgcg gcatcagagc agattgtact gagagtgcac catatgcggt gtgaaatacc 4260
gcacagatgc gtaaggagaa aataccgcat caggcgccat tcgccattca ggctgcgcaa 4320
ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgg cgaaaggggg 4380
atgtgctgca aggcgattaa gttgggtaac gccagggttt tcccagtcac gacgttgtaa 4440
aacgacggcc agtgccaagc ttgcatgcct gcaggtcgac gaccgagcgc gcgccaccca 4500
gcctatcccg cgcgggtcgg gacccaaaat aagcgggccc cgccgcgccc cgtcgggcga 4560
gcgggtgtat ctacgaacgg aactgggagg cgactcggaa gagtttggtt agaaagggga 4620
acaccatcgc ggacggccca gtgctctggd cagctgagcg tgcattgtgt tcaattctga 4680
cctgtggcat gtaaggaacg tgctcgggat cggaggggtgg cgcgagagcc tcttcggtgt 4740
gagattagta actgtactgc gaagccgcgg aggggttagg atgagaggta gacagggtcg 4800
cagcccaggt gcgagaagga ctgcgaagga ctgttcttcg accgcgcacc tgcaattgcg 4860
cgcatggata gaatagagcg tcgccctcga gggggactcg accagggctg gtggtggcgc 4920
ccgacgggac tggctgggca tttgcagatg gcgcgcagtc caggccgccg ccgatgtgtt 4980
catcccgttt tgtcagtatc gatcggatct ttcgggcgtg ggtataaaag cgcgccgccc 5040
gccgtctccc tctttctcca gcactcccat ccagagcact tccctctccc atcgcatccc 5100
atcacacaat aatgcccatc ac 5122

<210> 13
<211> 5490
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic vector

<400> 13
agcttctccg gccccgaatc gaacggcagg atgtgtgggc gtgtccaata ttgccatgaa 60
aatctgtcag aagtgagccc tctcgtcacc ctgtacagct tcgctgagtt gaaaagcagg 120
gttcatcttg ggctcactga tgcactgagc tcgaccggag aactaaatga ccagccggag 180
tgttcactaa cttaacgccg ggtattcagg gcagcttctc tatgttgcgc ctacgacgta 240
gatcaccgcc catgaacggg ggaaacgggg aggggtgcgt ttggtacgtc tttacgtctg 300
gctatgttgt attgaccagc gtctgcagaa gatgggcacg acgatgcgcc gagccggcca 360
gtgtcgtcgg atgtccactg ttgaggccat cctttttgcta gacagacgga agagctttgg 420
aggtgcgatt cctctacgaa tgggaagggg cttagatgga gagtgacacg tctgagctcc 480
ccaacacgcc ttcgccgagg gtgcgtctcc gcggacattc acctcagttc attgttctga 540
cctgcctaat tgtatagacc ggccaacaac cttgctgacg cccatcataa cagtgccctg 600
cacagagcct tcccactcag tcggcgcctc cctcaatcaa tcccactaac tcgccggctc 660
tgccccttcg ccgctcgaca cgtcgcttgg aagagcccgg gcacgggcgt ccgctccccc 720
cttccctccg cgtcgtcatg cacgcagcgt taatgttgct gcaggcgagc cgtaagtata 780
ttcaaaggcg tagcgaatga atagcaggcg cgcggggacc tggcacgcgc ggcatgaaca 840
tgcagacttg ggtgacgata acttgaactc agacgcggcg aatgaatatc caaacgcgcg 900
ggaagaaaat aatttacggg agcctcccca ggtataaaag cccctcaccc gctcactctt 960
tctccagtcg aacaccccag ttcaactacc cagccccttcc ttccttcgct atccttcytt 1020
acaacctgct cgccatggga tatcgcatgc ctgcagagct ctagactcga cgggcccggt 1080
accgcggccg ccttaagacg cgtggatccg caggtgaacg cgcctatcgg tgggatattc 1140
gggcgacggg agcctcggca atctgagcct cgttactgcc tagcaaattc ggaatccctt 1200
cgatgtcata gggtcgcgga caagtgatcg tcttgctaca tactccaagg tgttgactca 1260
ttccctcgat aatgaacatt gttgttgttg tttgttctct atccgctcag tcacgcgacc 1320
ccacacgtgc atggttgaac ttcgccacgc aacaaccgca tgacgacatg gcgaacctaa 1380
gtaaaggctg agtcgtggac taaagcactc cactttacgg cgaggatgcc agtctacgtc 1440
atgaatgaag cctcaggtcc cgaagtaagg gggtacaaaa ggagggtgaa aggtggacgt 1500

```
tttcttacca tccttccacc tcccagacca ccatgccggg aattcccagc ttgctcaaaa 1560
aggttctgcc cgtacgcccg cgaaattcct tcgaggtggc ccctatcgca tacatgcacg 1620
acttcaaaac atccattcta tcattttggg atcgtacaat tattagacat gttgtacaac 1680
gttacattcc tttcttcttt tactctccgg cccagtctat gtagaggtaa agtacaagcg 1740
tccaaaggat caggcactta gagcgcgccg tcttgcttcg ccgcttagag cgcgccgtcc 1800
tgcttcgccg cgtagacgag caggtcgcag acacggcggg agtagcccca ctcgttgtcg 1860
taccaggcaa tgagcttcac gaagctcttg ctgatcgcga tgccggggat cgatccacgc 1920
gtcttaaggc ggccgcggta ccccctcgga cccgtcgggc cgcgtcggac cggcggtgtt 1980
ggtcggcgtc ggtcagtcct gctcctcggc cacgaagtgc acgcagttgc cggccgggtc 2040
gcgcagggcg aactcccgcc cccacggctg ctcgccgatc tcggtcatgg ccggcccgga 2100
ggcgtcccgg aagttcgtgg acacgacctc cgaccactcg gcgtacagct cgtccaggcc 2160
gcgcacccac acccaggcca gggtgttgtc cggcaccacc tggtcctgga ccgcgctgat 2220
gaacagggtc acgtcgtccc ggaccacacc ggcgaagtcg tcctccacga agtcccggga 2280
gaacccgagc cggtcggtcc agaactcgac cgctccggcg acgtcgcgcg cggtgagcac 2340
cggaacggca ctggtcaact tggccatgca tggtgatggg cattatgtgt gatgggatgc 2400
gatgggagag ggaagtgctc tggatgggag tgctggagaa agagggagac ggcgggcggc 2460
gcgcctttta tacccacgcc cgaaagatcc gatcgatact gacaaaacgg gatgaacaca 2520
tcggcggcgg cctggactgc gcgccatctg caaatgccca gccagtcccg tcgggcgcca 2580
ccaccagccc tggtcgagtc cccctcgagg gcgacgctct attctatcca tgcgcgcaat 2640
tgcaggtgcg cggtcgaaga acagtccttc gcagtccttc tcgcacctgg gctgcgaccc 2700
tgtctacctc tcatcctaac ccctccgcgg cttcgcagta cagttactaa tctcacaccg 2760
aagaggctct cgcgccaccc tccgatcccg agcacgttcc ttacatgcca cagcgtcaga 2820
attgaacaca atgcacgtca ratcagatcc ccgggaattc gtaatcatgg tcatagctgt 2880
ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc ggaagcataa 2940
agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac 3000
tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 3060
cggggagagg cggtttgcgt attgggcgct cttccgcttc ctcgctcact gactcgctgc 3120
gctcggtcgt tcggctgcgg cgagcggtat cagctcactc aaaggcggta atacggttat 3180
ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca 3240
ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc 3300
atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc 3360
aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg 3420
gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta 3480
ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg 3540
ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggtaagac 3600
acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag 3660
gcggtgctac agagttcttg aagtggtggc ctaactacgg ctacactaga aggacagtat 3720
ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat 3780
ccggcaaaca aaccaccgct ggtagcggtg gttttttttgt ttgcaagcag cagattacgc 3840
gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt 3900
ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct 3960
agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt 4020
ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc 4080
gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg gagggcttac 4140
catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct ccagatttat 4200
cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca actttatccg 4260
cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg ccagttaata 4320
gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt gtcacgctcg tcgtttggta 4380
tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc cccatgttgt 4440
gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt cagaagtaag ttggccgcag 4500
tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg ccatccgtaa 4560
gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc 4620
gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat agcagaactt 4680
taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg atcttaccgc 4740
tgttgagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca gcatctttta 4800
ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa 4860
taagggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat tattgaagca 4920
tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac 4980
```

aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta 5040
ttatcatgac attaacctat aaaaatаggc gtatcacgag gccctttcgt ctcgcgcgtt 5100
tcggtgatga cggtgaaaac ctctgacaca tgcagctccc ggagacggtc acagcttgtc 5160
tgtaagcgga tgccgggagc agacaagccc gtcaggcgcc gtcagcgggt gttggcgggt 5220
gtcggggctg gcttaactat gcggcatcag agcagattgt actgagagtg caccatatgc 5280
ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggcgc cattcgccat 5340
tcaggctgcg caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc 5400
tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt 5460
cacgacgttg taaaacgacg gccagtgcca 5490

<210> 14
<211> 6983
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic vector

<400> 14
catgggatat cgcatgcctg cagagctcta gagtcgacgg gcccggtacc gcggccgcct 60
taagacgcgt ggatccgcag gtgaacgcgc ctatcggtgg gatattcggg cgacgggagc 120
ctcggcaatc tgagcctcgt tactgcctag caaattcgga atcccttcga tgtcataggg 180
tcgcggacaa gtgatcgtct tgctacatac tccaaggtgt tgactcattc cctcgataat 240
gaacattgtt gttgttgttt gttctctatc cgctcagtca cgcgacccca cacgtgcatg 300
gttgaacttc gccacgcaac aaccgcatga cgacatggcg aacctaagta aaggctgagt 360
cgtggactaa agcactccac tttacggcga ggatgccagt ctacgtcatg aatgaagcct 420
caggtcccga agtaaggggg tacaaaagga gggtgaaagg tggacgtttt cttaccatcc 480
ttccacctcc cagaccacca tgccgggaat tcccagcttg ctcaaaaagg ttctgcccgt 540
acgcccgcga aattccttcg aggtggcccc tatcgcatac atgcacgact tcaaaacatc 600
cattctatca ttttgggatc gtacaattat tagacatgtt gtacaacgtt acattccttt 660
cttcttttac tctccggccc agtctatgta gaggtaaagt acaagcgtcc aaaggatcag 720
gcacttagag cgcgccgtct tgcttcgccg cttagagcgc gccgtcctgc ttcgccgcgt 780
agacgagcag gtcgcagaca cggcgggagt agccccactc gttgtcgtac caggcaatga 840
gcttcacgaa gctcttgctg atcgcgatgc cggggatcga tccacgcgtc ttaaggcggc 900
cgcggtaccc cctcggaccc gtcgggccgc gtcggaccgg cggtgttggt cggcgtcggt 960
cagtcctgct cctcggccac gaagtgcacg cagttgccgg ccgggtcgcg cagggcgaac 1020
tcccgccccc acggctgctc gccgatctcg gtcatggccg gcccggaggc gtcccggaag 1080
ttcgtggaca cgacctccga ccactcggcg tacagctcgt ccaggccgcg cacccacacc 1140
caggccaggg tgttgtccgg caccacctgg tcctggaccg cgctgatgaa cagggtcacg 1200
tcgtcccgga ccacaccggc gaagtcgtcc tccacgaagt cccgggagaa cccgagccgg 1260
tcggtccaga actcgaccgc tccggcgacg tcgcgcgcgg tgagcaccgg aacggcactg 1320
gtcaacttgg ccatgcatgg tgatgggcat tatgtgtgat gggatgcgat gggagaggga 1380
agtgctctgg atgggagtgc tggagaaaga gggagacggc gggcggcgcg ccttttatac 1440
ccacgcccga aagatccgat cgatactgac aaaacgggat gaacacatcg gcggcggcct 1500
ggactgcgcg ccatctgcaa atgcccagcc agtcccgtcg ggcgccacca ccagccctgg 1560
tcgagtcccc ctcgagggcg acgctctatt ctatccatgc gcgcaattgc aggtgcgcgg 1620
tcgaagaaca gtccttcgca gtccttctcg cacctgggct gcgaccctgt ctacctctca 1680
tcctaacccc tccgcggctt cgcagtacag ttactaatct cacaccgaag aggctctcgc 1740
gccaccctcc gatcccgagc acgttcctta catgccacag cgtcagaatt gaacacaatg 1800
cacgtcarat cagatccccg ggaattcgta atcatggtca tagctgtttc ctgtgtgaaa 1860
ttgttatccg ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg 1920
gggtgcctaa tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca 1980
gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg 2040
tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg 2100
gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg 2160

ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 2220
ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg 2280
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc 2340
tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc 2400
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc 2460
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg 2520
ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc 2580
actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 2640
gttcttgaag tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc 2700
tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac 2760
caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 2820
atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 2880
acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa 2940
ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta 3000
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 3060
tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag 3120
tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag caataaacca 3180
gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc 3240
tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt 3300
tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag 3360
ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt 3420
tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat 3480
ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt 3540
gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc 3600
ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat 3660
cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag 3720
ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt 3780
ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg 3840
gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta 3900
ttgtctcatg agcggatACa tatttgaatg tatttagaaa aataaacaaa taggggttcc 3960
gcgcacattt ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt 4020
aacctataaa aataggcgta tcacgaggcc ctttcgtctc gcgcgtttcg gtgatgacgg 4080
tgaaaacctc tgacacatgc agctcccgga gacggtcaca gcttgtctgt aagcggatgc 4140
cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcgggtgtc ggggctggct 4200
taactatgcg gcatcagagc agattgtact gagagtgcac catatgcggt gtgaaatacc 4260
gcacagatgc gtaaggagaa aataccgcat caggcgccat tcgccattca ggctgcgcaa 4320
ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgg cgaaaggggg 4380
atgtgctgca aggcgattaa gttgggtaac gccagggttt tcccagtcac gacgttgtaa 4440
aacgacggcc agtgccaagc ttagatctcc gaaccagaaa tgcgattgcg ttcaggccca 4500
attaagaata aagctgcgtc agggcagcga cgtatcttga tccatcattg actcaccggc 4560
atcggcgtca acaccaaagc aagctcgtcc cacccatagg cgtgcaccgg ccggcgtgcg 4620
ccattgaggt acatgagcgg ggcgaaagtc cgccattggt agccctgtcg tggacgcgcg 4680
gcgatgaaac gtttcccacc attgggaaga aacgtctgcg gcccatcatc ccttcaccgg 4740
atgacaaggc ggcgtcgcgc ctttgccgca gaggccggcg ggcgacatgc acagcgaagg 4800
tccgttgcgg atgggaagca ggcaatcagt gggtgtccta cgccgccacg atggtcgggg 4860
agcgtaggcg ccctcccata aggcggcaag catcatgatg ctctccgatt cgggaagcct 4920
ggtgcgatgc tggagagact ctctccgaga gaccagtgtg cgcaacgttc ctggcctgga 4980
agactttaaa gtgagtgtag aagggcgagc agaggacgat catcggattg caggaaccat 5040
cggcatcctc agcctgggaa ggatggctct tggtagacat tcgcggaagg tgtcctagat 5100
gtgagcgggc ttcttggatg atcatgtcgt aacttttttct gacctcgtcg gtggtacgca 5160
tggcaggatt gagcattacg gtatgcctcc cattcataaa cgataacccc ttccttcagg 5220
ttggtcatct ccatagagcg gcacgctctc aaggcctagg ctattcacac ctccttcgca 5280
acatccctat tcacggtgtc tgtaaggaac gacttgtcat gggatcacat gaagtgcagc 5340
atactgttcg ccggtctcgc agtacagacg ctagtacggg aagtcgacat ccaagcgttc 5400
agtcaccaca tggcaaaaaa gctgcaccat actctttatg gtgagttgtt cgtgagtggt 5460
atacagtcat tcatgaggga atgcccaccg gatagggtgt ggcggccgca atattcatcg 5520
cctggcaata gtcgatgtgc gtccttgttc aatgaatatc atgggtcaca tgtggagacg 5580
gttaaacagc gttgactgtg aatccctggt gtgtgttggg ccgaacaggt acgttgcagg 5640

aacaccaata tctcttcggc agcccagttc tttgcgagcg gcacaggcag gcatcgcgca 5700
acagatccca gccatccggc ctctgacatt cgggatacct gaagcccttc aggtacggag 5760
cgaagaggtg ggctctctgc agcgattggc ggacggatag ctgtatttcc tctctcacca 5820
ttgggaagat gtgaaaggct ccatcatata gcggctcaac tctacctcga atgtccaaac 5880
acggcgggaa tacttattta tgtggacaag gccgagctat gatagcttgc tcccgaagtt 5940
ggtaagtccc gcaatctgcg gttcaggcaa cagtctcgga aaaataagaa gaatattgta 6000
ggtgcgtgta ggcgtatcgc ccaaatgcgc acacacggag gctttaggag atgaagcgcc 6060
cgtgagcggt aagggagttg gttcaccgcc gccccgaccg actctctctc tttcccagca 6120
tcatgtctcg gcgcaaactt taccctctat tgaccaactc cacgagaaag caggaacagc 6180
ttccttgtct ctcatgacgt ccgcaatcca gacccttagc cggttcgtta ctcatcgtta 6240
tccctgccgc catcgtagtg gagtcagcct ggccagtgcg tagtcccgtc tctcttgctg 6300
cactagagaa gccccatgag acagcgtttt ttgcttttatt tctgctgttt ctatagacac 6360
catagggggca aacgatcctg cacgcccaga ggtattgggc tcgtcagatt cccagttttt 6420
ctcctcggtc tgaatcggct gcacggcaga taaatcggcc ggaaatgcta tagcccttca 6480
tagcccgcta tgagagtcgc aaaaggcttg tcagtcaggt cggtcgagtg gctctcacga 6540
agagcgtcaa cttcgcgcga cagccgcctt tcagggcaag atagatcctc ccatcatccc 6600
ctactgcgct cagcgccggt accgaacaat tgacttaccg acatcctccg ggacgcgcaa 6660
atgctgttcg acggaacgta atcctcttcg tcccgcctct tttcgctctc acgcattccg 6720
tgtggttcgc gcgacggccg ctcatcagga ccagaccagt ctcaatgtct ggtaccggca 6780
caatggtgac actgcggcaa ctgagtaggt ctggtcactc tggtgcaccg tcgcttacgc 6840
tgaccttcgg gatactgtcc tgcagacatc tggagcgcct gtctttcccc tagtataaat 6900
gatgtctgtc cgcaggtcct tgaagaccgc tcgagtccca cttgagtttt aggtaggacc 6960
tgttcctcca caacccctct ttc 6983

<210> 15
<211> 4395
<212> DNA
<213> Pycnoporus cinnabarinus

<400> 15
gagcttactg gatcttccag agaaatcgtt ggagaggtcg gccggtcagc ctcaccgaca 60
tttgacgtgg ccgatgattc tgtggatgcc atcgtgttgg atcctgagct tgcatccatc 120
gccaagcgcg tcaaagctga ggtgggaagg cagggaggca caccagttcc cgaaggaggc 180
ggacctgaga ttgtgacgct caagattata tggaaacctc atccgctgaa ccccaacggc 240
cgtccggaac tctgggctat gaagcagaga cgggtaggtg aagtcgctca tcacgcctcg 300
ttcttactca ccatcttctg cagcacgaga atttccaccg gctttgttcc gaagtagcgg 360
acctcgcgag tgttcgtagt gagaacgtcg tgcttttccct cgacgggaaa cgcgtgttcc 420
cttcatctac ccctcacagt gtcggtgtct gggcagaagc tgagctaggt tagtgactta 480
tcctgtgcgt gacggcacga tgcttactct tcaacagaag cttgtgacaa gatcacctac 540
caatacctgc aggaaaataa gcgaatgcgt tccgaatccg ttgctccgcc aacccatcct 600
cacctcgacg acatttcccg tcagtctcca actcgcgcgc gctccccttc catcaccgag 660
ctgtccgaga atgaatccgg cgctgcagag tctggtcctg aagataaggg cacaagcact 720
ggggaggcct tcagcctgat actcgtgagc gaacggacca agggcaagcg aataaccctc 780
cgtgtgctcc caaccaccaa atgtggcgtc atagttcgca agttcctaga gaaggccggc 840
ttacaggacg aataccccga tgtcaccctt gctgcgaatg ggcgaggacg cacgaagacg 900
tcagccaaga caccggcgct gagcgtcgat ggagataaga tggatccgga ggcacctatt 960
ggtgatgccg acctggaaga tggggatcaa gtcgaggtgg ttggtctttg atgtagcgag 1020
tgcgtggtgt tacgttttcg tcttgctatc ggtttttctg ctcttgtctt gttagtaagt 1080
agtataatga tggataatca cacaacaacg tatgtgttcc agggacttct ctctcagtgg 1140
gtgtgtggct gattgtacga aacatcgcac aggcctttca ccgctgctcc tagcgcgaga 1200
ccacatgaac gccctcgcac gtcagtcggc ctcgcgaacg atagggcagt gccaaatgca 1260
ggcgaaaatg actcagttag gccacgcctg cgctttaact ttagcgttct aataccctcg 1320
acctccgtag cggtcctgtg agcgcggaaa gcttggcatc tacgctcttt gtccggtccg 1380
aagccactac attgagcggt tcgccccggc ggtgacaaac tgcgaggcgc aagaatgtag 1440
ccgggcctgc ggaaaggtca tgaacaaact atgtcggccc caaccagtgc taccgacacc 1500
gttctccgtg tttcagtatg ccttcagctg tcggtgggcg gggtggctcc gatatgtgta 1560
ctcggaaacg ctcacagcgc tctttgtatt gccgggtatg tgaccaacgg tgccctcatg 1620

ctcttcgctt gttgatgctc cttcaggaca ccgtctggga ctctggcaag tcagctgctg 1680
ctcgccacag ttctaggaac gtctcaatgc tcctaggcgg cggttacagc aaacgccttg 1740
caccgggatg ggcctcggta cgccgcacag gcgaggctgt cctactcggc gttcgttagt 1800
agccccccat ccacgtaaga gtacctcctg cagccaccat cgtctactag cgtaccaccc 1860
acgtccactc acatcatatg ccgcccgacg cccccggact gattccgcgc tattgttgag 1920
atataagagg agtgttcgaa cggaccaagg agccataatc ccctcgagca tttcgagatc 1980
ctctccccac tgaactcctt cgccgtcacc acaaaacctc gcgtagatgt cacacttcat 2040
cgttactggg cctgtaggag gtcagactga gggcgctcct gctcccaacc gcctcgaaat 2100
caacgacttc gtcaagaatg aggagttctt ctcgctttac gtccaggctc tcggtgcgtc 2160
gccttggcac atgtatgctc acccctatta ccatgaagct catgagccct cactacatac 2220
agatatcatg tatggactga agcaggagga actgatctcg ttcttccaga tcggtggcat 2280
tcatggattg ccatacgttg cctggagtga tgccggagcg gatgaccctg ctgagccgtc 2340
cgggtactgt acccatggct ccgtactgtt cccgacctgg cataggcctt acgtcgcact 2400
atatgaggta agcagcttgc tagatcagac cgctacggac gacgctgaga ctcaaaatgg 2460
ctacagcaaa tcttgcacaa gtatgctgga gagatcgctg ataagtacac ggtcgacaaa 2520
ccgcgttggc agaaggcagc ggccgacctg cgccaaccct tctgggactg ggccaagaac 2580
acgctgcctc ctcctgaagt catctctctc gacaaagtca cgattacgac accagatgga 2640
cagaggacgc aagttgacaa tccactccgt cgctaccgct tccatccgat cgaccccagc 2700
ttcccagagc catacagcaa ctggccagcg acactgagac atccgacaag tgatggctcg 2760
gatgccaaag acaacgtgaa ggatctcact acgtaagcca attcgccata aagacgctcc 2820
tccattcatc tcaatgtata tatgtgacag tactctgaag gcggaccagc ctgatatcac 2880
gacgaagacg tataatctat tgaccagagt gcacacgtgg ccggcgttca gcaaccacac 2940
tccaggcgat ggcggcagct ccagtaacag tcttgaggcc attcacgacc acatccatga 3000
ctcagttggc ggcggaggcc agatgggaga cccgtccgtg gcaggtatgt gaagtgattc 3060
ttcgcgagag acgtgactta catgtccttg taggcttcga cccaatcttc ttcctgcacc 3120
attgccaagt tgatcgtctt cttgcactgt ggtccgcctt gaaccccggc gtgtgggtca 3180
acagctctag ctccgaagat ggcacctaca cgatcccgcc tgactctacc gtggaccaaa 3240
ctactggtgg gttcccgcac agctgtgcgc tgtggagtcg ccgttgactt ccatcactct 3300
cagcattgac gcccttctgg gatacccaaa gcacattctg gacgtccttc cagtctgctg 3360
gagtctcgcc cagccaattt ggctattctt accccgagtt taacggtctc aacctgcaag 3420
atcagaaggc tgtgaaagat cacatcgccg aggtcgtgaa cgagctctac ggtcatcgca 3480
tgcggaaaac cttccctttc ccccagctcc aggcagtttc cgtagccaag cagggcgacg 3540
ccgtcactcc atccgtggct accgattcag tgtcgtcttc taccacacct gccgaaaatc 3600
ccgcatcccg cgaggatgcc tctgataagg acacagagcc gacgctcaat gtagaggttg 3660
ccgcgccagg cgcgcacttg acctccacca agtattggga ctggactgct cgcattcatg 3720
tcaagaagta cgaagtcgga ggcagcttca gcgtcctgct cttcctgggt gcaatccccg 3780
agaacccagc ggattggcgc acgagcccca actacgttgg cggtcatcat gctttcgtga 3840
atagctcacc gcagcgctgc gctaactgcc gtggtcaagg cgaccttgtc atcgaaggct 3900
tcgtccatct caacgaggcg atcgcccgcc atgcgcacct cgactccttc gatccaaccg 3960
tcgtgaggcc gtacctcacg cgcgagttgc actggggtgt gatgaaggtc agtgcctaca 4020
ctctgcatac gaccgtatat gtcgctaatt agatctatca aggtgaatgg caccgtcgtg 4080
cccctgcaag acgtcccgtc gctcgaggtt gtcgtcctct caactcctct tacccttcct 4140
ccgggagagc cattccctgt ccccggaacg cccgtcaatc atcatgacat cacccatgga 4200
cgtcctggtg gctctcacca cacgcactaa gcatgctgat ggcctgcccc tattgattaa 4260
acacgagtcg acctgagaac acatacaatg gatgtaatca tacttcactt ttgatgacaa 4320
tcgcttccac attctgttcc tagcgggaca gataacccag tcaaaaaaaa aaaaaaaaaa 4380
aaaacactgt catgc 4395

<210> 16
<211> 618
<212> PRT
<213> Pycnoporus cinnabarinus

<400> 16
Met Ser His Phe Ile Val Thr Gly Pro Val Gly Gly Gln Thr Glu Gly
1 5 10 15

```
Ala Pro Ala Pro Asn Arg Leu Glu Ile Asn Asp Phe Val Lys Asn Glu
                20                  25                  30

Glu Phe Phe Ser Leu Tyr Val Gln Ala Leu Asp Ile Met Tyr Gly Leu
            35                  40                  45

Lys Gln Glu Glu Leu Ile Ser Phe Phe Gln Ile Gly Gly Ile His Gly
        50                  55                  60

Leu Pro Tyr Val Ala Trp Ser Asp Ala Gly Ala Asp Asp Pro Ala Glu
65                  70                  75                  80

Pro Ser Gly Tyr Cys Thr His Gly Ser Val Leu Phe Pro Thr Trp His
                85                  90                  95

Arg Pro Tyr Val Ala Leu Tyr Glu Gln Ile Leu His Lys Tyr Ala Gly
            100                 105                 110

Glu Ile Ala Asp Lys Tyr Thr Val Asp Lys Pro Arg Trp Gln Lys Ala
            115                 120                 125

Ala Ala Asp Leu Arg Gln Pro Phe Trp Asp Trp Ala Lys Asn Thr Leu
    130                 135                 140

Pro Pro Pro Glu Val Ile Ser Leu Asp Lys Val Thr Ile Thr Thr Pro
145                 150                 155                 160

Asp Gly Gln Arg Thr Gln Val Asp Asn Pro Leu Arg Arg Tyr Arg Phe
                165                 170                 175

His Pro Ile Asp Pro Ser Phe Pro Glu Pro Tyr Ser Asn Trp Pro Ala
            180                 185                 190

Thr Leu Arg His Pro Thr Ser Asp Gly Ser Asp Ala Lys Asp Asn Val
            195                 200                 205

Lys Asp Leu Thr Thr Thr Leu Lys Ala Asp Gln Pro Asp Ile Thr Thr
    210                 215                 220

Lys Thr Tyr Asn Leu Leu Thr Arg Val His Thr Trp Pro Ala Phe Ser
225                 230                 235                 240

Asn His Thr Pro Gly Asp Gly Gly Ser Ser Ser Asn Ser Leu Glu Ala
                245                 250                 255

Ile His Asp His Ile His Asp Ser Val Gly Gly Gly Gly Gln Met Gly
            260                 265                 270

Asp Pro Ser Val Ala Gly Phe Asp Pro Ile Phe Phe Leu His His Cys
            275                 280                 285

Gln Val Asp Arg Leu Leu Ala Leu Trp Ser Ala Leu Asn Pro Gly Val
    290                 295                 300

Trp Val Asn Ser Ser Ser Ser Glu Asp Gly Thr Tyr Thr Ile Pro Pro
305                 310                 315                 320
```

Asp Ser Thr Val Asp Gln Thr Thr Ala Leu Thr Pro Phe Trp Asp Thr
325 330 335

Gln Ser Thr Phe Trp Thr Ser Phe Gln Ser Ala Gly Val Ser Pro Ser
340 345 350

Gln Phe Gly Tyr Ser Tyr Pro Glu Phe Asn Gly Leu Asn Leu Gln Asp
355 360 365

Gln Lys Ala Val Lys Asp His Ile Ala Glu Val Val Asn Glu Leu Tyr
370 375 380

Gly His Arg Met Arg Lys Thr Phe Pro Phe Pro Gln Leu Gln Ala Val
385 390 395 400

Ser Val Ala Lys Gln Gly Asp Ala Val Thr Pro Ser Val Ala Thr Asp
405 410 415

Ser Val Ser Ser Ser Thr Thr Pro Ala Glu Asn Pro Ala Ser Arg Glu
420 425 430

Asp Ala Ser Asp Lys Asp Thr Glu Pro Thr Leu Asn Val Glu Val Ala
435 440 445

Ala Pro Gly Ala His Leu Thr Ser Thr Lys Tyr Trp Asp Trp Thr Ala
450 455 460

Arg Ile His Val Lys Lys Tyr Glu Val Gly Gly Ser Phe Ser Val Leu
465 470 475 480

Leu Phe Leu Gly Ala Ile Pro Glu Asn Pro Ala Asp Trp Arg Thr Ser
485 490 495

Pro Asn Tyr Val Gly Gly His His Ala Phe Val Asn Ser Ser Pro Gln
500 505 510

Arg Cys Ala Asn Cys Arg Gly Gln Gly Asp Leu Val Ile Glu Gly Phe
515 520 525

Val His Leu Asn Glu Ala Ile Ala Arg His Ala His Leu Asp Ser Phe
530 535 540

Asp Pro Thr Val Val Arg Pro Tyr Leu Thr Arg Glu Leu His Trp Gly
545 550 555 560

Val Met Lys Val Asn Gly Thr Val Val Pro Leu Gln Asp Val Pro Ser
565 570 575

Leu Glu Val Val Val Leu Ser Thr Pro Leu Thr Leu Pro Pro Gly Glu
580 585 590

Pro Phe Pro Val Pro Gly Thr Pro Val Asn His His Asp Ile Thr His
595 600 605

Gly Arg Pro Gly Gly Ser His His Thr His
610 615

<210> 17
<211> 3037
<212> DNA
<213> Halocyphina villosa

<220>
<221> CDS
<222> (260)..(388)

<220>
<221> CDS
<222> (444)..(494)

<220>
<221> CDS
<222> (559)..(663)

<220>
<221> CDS
<222> (713)..(727)

<220>
<221> CDS
<222> (779)..(802)

<220>
<221> CDS
<222> (850)..(975)

<220>
<221> CDS
<222> (1024)..(1134)

<220>
<221> CDS
<222> (1188)..(1286)

<220>
<221> CDS
<222> (1333)..(1434)

<220>
<221> CDS
<222> (1515)..(1575)

<220>
<221> CDS
<222> (1625)..(1845)

<220>
<221> CDS
<222> (1973)..(2140)

<220>
<221> CDS
<222> (2187)..(2201)

<220>
<221> CDS
<222> (2249)..(2311)

<220>
<221> CDS
<222> (2367)..(2414)

<220>
<221> CDS
<222> (2463)..(2489)

<220>
<221> CDS
<222> (2542)..(2646)

<220>
<221> CDS
<222> (2697)..(2834)

<400> 17
tggggagatg gttctatata tcaaaatgat cttctgtcct gagctttcct cgtccttgtt 60

ttcgtcttgt cagtgccgcg acatgctttt attaaaccat tggcgagctg cccgcgccca 120

aggagatagc ataatcgcct gagaaaccta gtcgtctcat ggccgtgtaa ccgttcttgc 180

gacttatttt cgcacttctc tcagaatata aaggcctatt gtgatacggt tcatctaacc 240

ccagcgtccc ctccgaaag atg ggc tgc ctc tca ctc ttc gca ttc ctt act 292
Met Gly Cys Leu Ser Leu Phe Ala Phe Leu Thr
1 5 10

gct tta aac tca gtt cat gcc gct gtg ggt ccc gtt acg gac tta aca 340
Ala Leu Asn Ser Val His Ala Ala Val Gly Pro Val Thr Asp Leu Thr
15 20 25

ctg atc gta gat act gtc gcc ccc gac ggt gct gct ttc gcg cgg gaa 388
Leu Ile Val Asp Thr Val Ala Pro Asp Gly Ala Ala Phe Ala Arg Glu
30 35 40

ggtgagactt tgcgactgta aatgccggat ttgagtttct aattataatc ttcca gcc 446
Ala

att gtc gtc caa gag gaa cca aac tcc gtc att ggt ccg gtc atc gta 494
Ile Val Val Gln Glu Glu Pro Asn Ser Val Ile Gly Pro Val Ile Val
45 50 55 60

ggtgggtagc tacgagtctt cctccttcat ttagctcatc accaagtgat atgatattaa 554

ttaa ggt caa aag ggg gac aac ttt cgg ctc aat gtt atc aac aat ttg 603
Gly Gln Lys Gly Asp Asn Phe Arg Leu Asn Val Ile Asn Asn Leu
65 70 75

gat tct ccg aac atg cgc caa tct act tcc att cat tgg cat ggc atc 651
Asp Ser Pro Asn Met Arg Gln Ser Thr Ser Ile His Trp His Gly Ile
80 85 90

```
ttc caa gga aac ggtacgtggt atatcggata  atctatctgt atccattgac          703
Phe Gln Gly Asn
           95

tcgaatata ggt cag aat tgg gct ggtgcgttgg ccttcctgaa gcctgctcga       757
          Gly Gln Asn Trp Ala
                      100

atttatcttc ctgaattttt a gat ggc gcc gca ttc gtt aac cag               802
                       Asp Gly Ala Ala Phe Val Asn Gln
                                   105

gtaaggagat gttcctgcct tcgtttcccc agaactaatt atcctag tgc ccc att       858
                                                  Cys Pro Ile
                                                      110

gcc ccc gga ggg gac tcg ttc ttg tac gac ttt acc gaa cct ttc cag       906
Ala Pro Gly Gly Asp Ser Phe Leu Tyr Asp Phe Thr Glu Pro Phe Gln
           115                 120                 125

act ggc aca ttt tgg tat cat tcc cat tta tca act caa tac tgc gat       954
Thr Gly Thr Phe Trp Tyr His Ser His Leu Ser Thr Gln Tyr Cys Asp
       130                 135                 140

gga ctg agg gga gca ttc gtc gttcgttctc ttcttcatca agtcaccgct         1005
Gly Leu Arg Gly Ala Phe Val
    145                 150

ttcttctcac ttatctag atc tac gat ccg ctc gac cct tac cgg ttg ctc      1056
                    Ile Tyr Asp Pro Leu Asp Pro Tyr Arg Leu Leu
                                155                 160

tac gat gtc gac gac gag tcg act gtg att act ctg gcg gac tgg tac      1104
Tyr Asp Val Asp Asp Glu Ser Thr Val Ile Thr Leu Ala Asp Trp Tyr
           165                 170                 175

cac agc tat gcg gag gac att cta atc gcg taggagattt tcccaagatg       1154
His Ser Tyr Ala Glu Asp Ile Leu Ile Ala
       180                 185

tctcctctgc ctctctgaaa tccatgaact agt gca ggc gac act atc ctc atc     1208
                                    Ala Gly Asp Thr Ile Leu Ile
                                            190

aat ggt cac gga aga ttc gcc gga gcc ggc gga acg gca aca gaa cta      1256
Asn Gly His Gly Arg Phe Ala Gly Ala Gly Gly Thr Ala Thr Glu Leu
195                 200                 205                 210

tct gtc att act gtt gag cat gga aag cgg taggcattct ccctcggctt       1306
Ser Val Ile Thr Val Glu His Gly Lys Arg
                215                 220

tgtagatgtg tctaatttgt gatagc tac cga ttg cga ttt gcc aat atc gct    1359
                           Tyr Arg Leu Arg Phe Ala Asn Ile Ala
                                       225
```

```
tgt gac cct tgg ttt gcc gtg aaa atc gat agc cat acg aac ctt cgc      1407
Cys Asp Pro Trp Phe Ala Val Lys Ile Asp Ser His Thr Asn Leu Arg
230                 235                 240                 245

gtt atc gaa gct gac ggt att act act gtgcctgtca cggtggactc            1454
Val Ile Glu Ala Asp Gly Ile Thr Thr
                250

cttcaatgta ggcttaccct tagcactttc ccactctgga tcctcttatg acttcccaag    1514

atc ttt gtg ggc caa cga tat agt gtc atc ctc cat gcc aac cag cct      1562
Ile Phe Val Gly Gln Arg Tyr Ser Val Ile Leu His Ala Asn Gln Pro
255                 260                 265                 270

gtt gga aac tac t gtaagctgcc taaatgttgc atgactgtcc atgattctaa        1615
Val Gly Asn Tyr


ccccgccag gg  att cgg gcc gct ccg aac ggc gtg agc aat ttc gcg ggt    1665
          Trp Ile Arg Ala Ala Pro Asn Gly Val Ser Asn Phe Ala Gly
          275                 280                 285

ggg atc gac tcg gct att ctc cgt tat gtt ggc gcc cca gaa gaa gag      1713
Gly Ile Asp Ser Ala Ile Leu Arg Tyr Val Gly Ala Pro Glu Glu Glu
    290                 295                 300

ccc aac act agt gag gat act cca tcc gac aca ctt caa gag cag gat      1761
Pro Asn Thr Ser Glu Asp Thr Pro Ser Asp Thr Leu Gln Glu Gln Asp
305                 310                 315                 320

ctt cac ccg ctg atc cta ccc ggc gcg cca ggc atc cac tcc cgt ggg      1809
Leu His Pro Leu Ile Leu Pro Gly Ala Pro Gly Ile His Ser Arg Gly
                325                 330                 335

gcc gcc gac gtt gtc cac acc gta tca atg gag ttt gtgagtgtgg           1855
Ala Ala Asp Val Val His Thr Val Ser Met Glu Phe
            340                 345

cgacttttct ggcccccttt attaatataa tctggttagg atggcgcaaa cttccaattc    1915

ctcctggatg gcgtggcctt ccagccgtgc gtcatctctt tcaaagaatt tatctag       1972

ctg acg att ttg aaa tgt agc ccg acc atg ccc gtc ctt ctg caa ata      2020
Leu Thr Ile Leu Lys Cys Ser Pro Thr Met Pro Val Leu Leu Gln Ile
    350                 355                 360

tta tcg gga gcg cag act gct aat acc ctt ctc ccg gcg gga tcc ttt      2068
Leu Ser Gly Ala Gln Thr Ala Asn Thr Leu Leu Pro Ala Gly Ser Phe
365                 370                 375                 380

atc caa gcg tcg cac aat gac atc gtg gag ctc aat ttc cca gct gtc      2116
Ile Gln Ala Ser His Asn Asp Ile Val Glu Leu Asn Phe Pro Ala Val
                385                 390                 395

aac gta gcc gct gtc ggt gga ccg tgcgtcccat ctttccttgc cagcttgaaa     2170
Asn Val Ala Ala Val Gly Gly Pro
            400
```

```
tttacgctct tttaga cat cca atc cat ctg tgagcgcagc gggacctttg          2221
                 His Pro Ile His Leu
                 405

gcttatggca tatgacttat tattagc cat ggc cat gca ttc gac gtt ata cgc    2275
                              His Gly His Ala Phe Asp Val Ile Arg
                              410                 415

tct gct gga acg aac tcc gat aac tgg ttc aat ccg gtattttcat           2321
Ser Ala Gly Thr Asn Ser Asp Asn Trp Phe Asn Pro
    420                 425                 430

tcgacttcca taagatgacg atggctcact atggttttta cccag cct cgc aga gat    2378
                                                Pro Arg Arg Asp


gtc gta tcc acc ggt acc gat cct aat gac aat gtg tacgtgtttc           2424
Val Val Ser Thr Gly Thr Asp Pro Asn Asp Asn Val
435                 440                 445

gctattgatt gtccgttttg atttgactgt tggacagc acc att cgc ttc cgg gcc    2480
                                        Thr Ile Arg Phe Arg Ala
                                                450

gac aac ccg tacgtaaact gctgaatctc tcgttgtctt tggttctcat              2529
Asp Asn Pro
        455

aatctcatca ga ggt cca tgg ttc ctt cac tgc cac att gac tgg cac ctt    2580
             Gly Pro Trp Phe Leu His Cys His Ile Asp Trp His Leu
                          460                 465

gaa ctc ggc ttt gct ttg gtg att gca gaa gcg cct agc gaa tgg gac      2628
Glu Leu Gly Phe Ala Leu Val Ile Ala Glu Ala Pro Ser Glu Trp Asp
    470                 475                 480

agc gac att aac cct cct ggtgcgctgc ctgtgaacct tttctcccta            2676
Ser Asp Ile Asn Pro Pro
485                 490

cacttgctaa gatcgctcta gct gcg tgg gat gac cta tgc cct acg ttc gct    2729
                     Ala Ala Trp Asp Asp Leu Cys Pro Thr Phe Ala
                                     495                 500

tgg ctt ctc ttt tac tat ttc aag ttt cct cac att ctc aac ttc aca      2777
Trp Leu Leu Phe Tyr Tyr Phe Lys Phe Pro His Ile Leu Asn Phe Thr
          505                 510                 515

gat atg atg ccc tgc cgc ctg agc agc agt aat cga gtt aag aac ctc      2825
Asp Met Met Pro Cys Arg Leu Ser Ser Ser Asn Arg Val Lys Asn Leu
        520                 525                 530

aac gtt gac taaggaaaaa gcaaagcaga atatgaaact ctcatttatc              2874
Asn Val Asp
    535

tttatatcga cacattcact attcaaccta cggattttcc ctcgcacctg aatttcggtg    2934
```

ctagatcccc atccttggtg gagtaggaaa gaaatttctt gtataaaacc catgggttct 2994

tctaccaata tatacataac gtccgtgggg ttagttaatt cgt 3037

<210> 18
<211> 536
<212> PRT
<213> Halocyphina villosa

<400> 18

```
Met Gly Cys Leu Ser Leu Phe Ala Phe Leu Thr Ala Leu Asn Ser Val
1               5                  10                  15

His Ala Ala Val Gly Pro Val Thr Asp Leu Thr Leu Ile Val Asp Thr
            20                  25                  30

Val Ala Pro Asp Gly Ala Ala Phe Ala Arg Glu Ala Ile Val Val Gln
        35                  40                  45

Glu Glu Pro Asn Ser Val Ile Gly Pro Val Ile Val Gly Gln Lys Gly
    50                  55                  60

Asp Asn Phe Arg Leu Asn Val Ile Asn Asn Leu Asp Ser Pro Asn Met
65                  70                  75                  80

Arg Gln Ser Thr Ser Ile His Trp His Gly Ile Phe Gln Gly Asn Gly
                85                  90                  95

Gln Asn Trp Ala Asp Gly Ala Ala Phe Val Asn Gln Cys Pro Ile Ala
            100                 105                 110

Pro Gly Gly Asp Ser Phe Leu Tyr Asp Phe Thr Glu Pro Phe Gln Thr
        115                 120                 125

Gly Thr Phe Trp Tyr His Ser His Leu Ser Thr Gln Tyr Cys Asp Gly
    130                 135                 140

Leu Arg Gly Ala Phe Val Ile Tyr Asp Pro Leu Asp Pro Tyr Arg Leu
145                 150                 155                 160

Leu Tyr Asp Val Asp Asp Glu Ser Thr Val Ile Thr Leu Ala Asp Trp
                165                 170                 175

Tyr His Ser Tyr Ala Glu Asp Ile Leu Ile Ala Ala Gly Asp Thr Ile
            180                 185                 190

Leu Ile Asn Gly His Gly Arg Phe Ala Gly Ala Gly Gly Thr Ala Thr
        195                 200                 205

Glu Leu Ser Val Ile Thr Val Glu His Gly Lys Arg Tyr Arg Leu Arg
    210                 215                 220

Phe Ala Asn Ile Ala Cys Asp Pro Trp Phe Ala Val Lys Ile Asp Ser
225                 230                 235                 240

His Thr Asn Leu Arg Val Ile Glu Ala Asp Gly Ile Thr Thr Ile Phe
                245                 250                 255
```

```
Val Gly Gln Arg Tyr Ser Val Ile Leu His Ala Asn Gln Pro Val Gly
            260                 265                 270

Asn Tyr Trp Ile Arg Ala Ala Pro Asn Gly Val Ser Asn Phe Ala Gly
        275                 280                 285

Gly Ile Asp Ser Ala Ile Leu Arg Tyr Val Gly Ala Pro Glu Glu Glu
    290                 295                 300

Pro Asn Thr Ser Glu Asp Thr Pro Ser Asp Thr Leu Gln Glu Gln Asp
305                 310                 315                 320

Leu His Pro Leu Ile Leu Pro Gly Ala Pro Gly Ile His Ser Arg Gly
                325                 330                 335

Ala Ala Asp Val Val His Thr Val Ser Met Glu Phe Leu Thr Ile Leu
            340                 345                 350

Lys Cys Ser Pro Thr Met Pro Val Leu Leu Gln Ile Leu Ser Gly Ala
        355                 360                 365

Gln Thr Ala Asn Thr Leu Leu Pro Ala Gly Ser Phe Ile Gln Ala Ser
    370                 375                 380

His Asn Asp Ile Val Glu Leu Asn Phe Pro Ala Val Asn Val Ala Ala
385                 390                 395                 400

Val Gly Gly Pro His Pro Ile His Leu His Gly His Ala Phe Asp Val
                405                 410                 415

Ile Arg Ser Ala Gly Thr Asn Ser Asp Asn Trp Phe Asn Pro Pro Arg
            420                 425                 430

Arg Asp Val Val Ser Thr Gly Thr Asp Pro Asn Asp Asn Val Thr Ile
        435                 440                 445

Arg Phe Arg Ala Asp Asn Pro Gly Pro Trp Phe Leu His Cys His Ile
    450                 455                 460

Asp Trp His Leu Glu Leu Gly Phe Ala Leu Val Ile Ala Glu Ala Pro
465                 470                 475                 480

Ser Glu Trp Asp Ser Asp Ile Asn Pro Pro Ala Ala Trp Asp Asp Leu
                485                 490                 495

Cys Pro Thr Phe Ala Trp Leu Leu Phe Tyr Tyr Phe Lys Phe Pro His
            500                 505                 510

Ile Leu Asn Phe Thr Asp Met Met Pro Cys Arg Leu Ser Ser Ser Asn
        515                 520                 525

Arg Val Lys Asn Leu Asn Val Asp
    530                 535
```